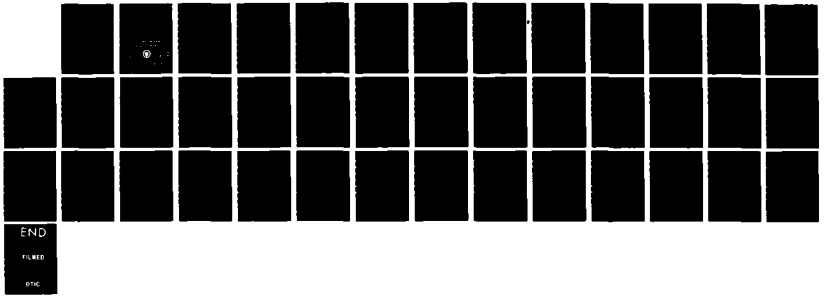
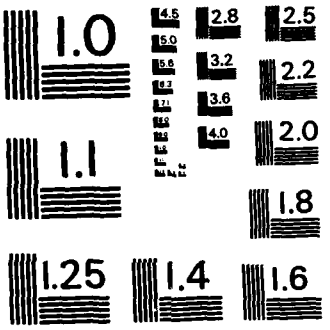
AD-A159 078 STRONG REPRESENTATION OF WEAK CONVERGENCE(U) PITTSBURGH 1/1
UNIV PA CENTER FOR MULTIVARIATE ANALYSIS
Z D BAI ET AL. JUL 85 TR-85-29 AFOSR-TR-85-0683
UNCLASSIFIED F49620-85-C-0008 F/G 12/1 NL



END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

REPRODUCED AT GOVERNMENT EXPENSE

AFOSR-TR- 85-0683

AD-A159 078

STRONG REPRESENTATION OF WEAK CONVERGENCE

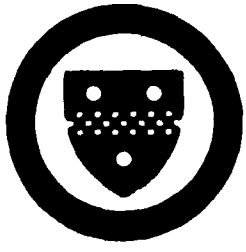
Z. D. Bai and W. Q. Liang

Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis

University of Pittsburgh

DTIC FILE COPY



DTIC
ELECTE
SEP 11 1985
S A D

85 9 10 117

Approved for public release,
distribution unlimited

STRONG REPRESENTATION OF WEAK CONVERGENCE

Z. D. Bai and W. Q. Liang

Center for Multivariate Analysis
University of Pittsburgh

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

July 1985

Technical Report No. 85-29

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

Research partially supported by the Air Force Office of Scientific Research (AFOSR) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

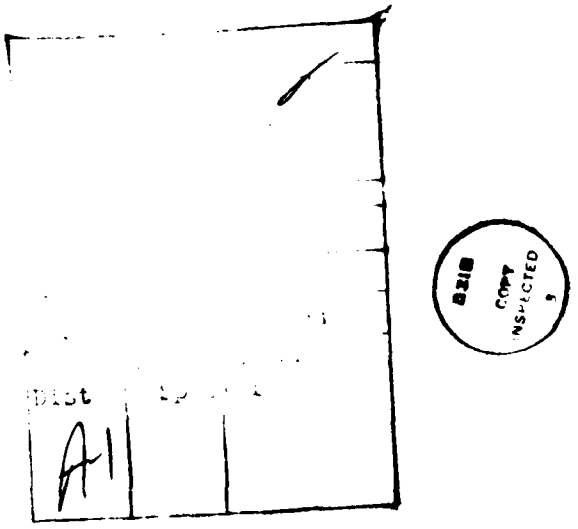
STRONG REPRESENTATION OF WEAK CONVERGENCE

Z. D. Bai and W. Q. Liang

ABSTRACT

Let μ_n, $n = 1,2,\ldots$, and μ be a given sequence of probability measures each of which is defined on a complete separable metric space S_n and S respectively. Also, a sequence of measurable mappings ϕ_n from S_n into S is given. In this paper, it is proved that if $\mu_n \circ \phi_n^{-1}$ weakly converge to μ, then there is a probability space (Ω,F,P), on which we can define a sequence of random elements X_n, from Ω into S_n, and a random element X, from Ω into S, such that μ_n is the distribution of X_n, μ is the distribution of X and $\lim_{n\to\infty} \phi_n(X_n) = X$ pointwise.

The result of Skorokhod (1956) is a special case of the result of this paper. Some applications in the area of random matrices, etc., are also given.



1. INTRODUCTION

It is well known that there is a big difference between the concepts of weak and strong convergence of random variables. In the area of limiting theory, it is of interest to study the difference as well as the link between the two concepts of convergence. Recent research work motivates us to investigate them. In Section 2, we shall prove the following theorem.

- Additional keywords: Skorokhod's theorem; finite dimension case; eigenvalues; random matrices.

THEOREM 1: Let S_n, $n = 1,2,\ldots,$ and S be complete separable metric spaces, with distance functions ρ_n and ρ respectively, and let ϕ_n be a measurable mapping from S_n into S. Suppose that μ_n and μ are probability measures defined on S_n and S, the Borel σ-fields deduced by the distances ρ_n and ρ, respectively, and suppose that $\mu_n \cdot \phi_n^{-1} \xrightarrow{W} \mu$. Then there is a probability space (Ω,F,P) and a sequence of S_n-valued random elements X_n, and an S-valued random element X, defined on (Ω,F,P), such that

1) X_n has distribution μ_n and X has distribution μ,

2) $\lim_{n\to\infty} \phi_n(X_n) = X$, pointwise.

In early 1956, Skorokhod proved a special case of Theorem 1, where $S_n = S$, for each n, and ϕ_n are all identity. It should be pointed out that our Theorem 1 is not a trivial generalization to Skorokhod's theorem. Theorem 1 played a key role in the proof of a theorem in Yin (1984), but Skorokhod's theorem is not applicable there.

Although Skorokhod's paper was published in early fifties, it seems that Skorokhod's theorem had not received much attention unfortunately. For instance, the Helley-Brary theorem can be easily obtained by Skorokhod theorem, but in

many recent probability textbooks, it was still proved by the approach of integration by parts. Even though the proof of Skorokhod's theorem seems a little complicated, we can give a very simple proof to the special case where $S_n = S = R^d$, the finite dimensional Euclidean space.

The power of Theorem 1 appears in the situation that we often encounter in large sample theory. Suppose that $\phi_n(Y_n) \xrightarrow{W} \phi$ and $F(\cdot,\cdot)$ is a two-variate continuous function. We are concerned with the limiting behavior of the roots of the equation $F(\phi_n(Y_n),X) = 0$. In general, the roots of $F(y,X) = 0$ do not have an obvious expression, but in many cases we can prove that the solution $x = x(y)$ is continuous in y. In these cases, by Theorem 1, we only need to investigate the behavior of the solution of $F(\phi,x) = 0$. Some concrete examples can be found in Bai (1984), Bai and Yin (1984) and Yin (1984).

We generalized Lusin's theorem to the measurable mapping from a complete separable metric space into another one. This result is stated in Theorem 2 and it played a key role in the proof of Theorem 1.

2. A GENERALIZATION OF SKOROKHOD'S THEOREM

We first assume that each ϕ_n is continuous. At the beginning, we construct a series of countable partitions of the space S as follows:

Let $B(x,r)$ denote the ball in S, with center x and radius r. Because S is separable, there is a countable set $\{x_i,\ i = 1,2,\ldots\}$, which is dense in S. Because there are at most countably many values of r such that $\mu(\partial B(x_i,r)) > 0$, for some i, where ∂B denotes the boundary of the set B. Thus for each k, there exists r_k, $2^{-(k+1)} < r_k < 2^{-k}$, being such that $\mu(\partial B(x_i,r_k)) = 0$ for any $i = 1,2,\ldots$. Write $C(k,1) = B(x_1,r_k)$, $C(k,i) = B(x_i,r_k) \setminus \bigcup_{j=1}^{i-1} B(x_j,r_k)$, and set

$$D_{i_1 i_2 \ldots i_k} = \bigcap_{j=1}^{k} C(j,i_j) \tag{1}$$

for any $i_1,i_2,\ldots,i_k = 1,2,\ldots$. It is obvious that $\{D_{i_1 \ldots i_k}\}$ satisfies the following properties:

$$\begin{aligned}
&1)\ D_{i_1 \ldots i_k} \cap D_{j_1 \ldots j_k} = \phi \quad \text{if} \quad (i_1,\ldots,i_k) \neq (j_1,\ldots,j_k),\\
&2)\ D_{i_1 \ldots i_{k-1}} = \bigcup_{i_k=1}^{\infty} D_{i_1 i_2,\ldots,i_k}, \qquad S = \bigcup_{i_1=1}^{\infty} D_{i_1},\\
&3)\ \mu(\partial D_{i_1 \ldots i_k}) = 0,\\
&4)\ d(D_{i_1 \ldots i_k}) < 2^{-k}, \text{ where } d(D) \text{ denotes the diameter of the set } D.
\end{aligned} \tag{2}$$

Using the same approach, for each n, we split S_n into partitions

$\{D^{(n)}_{i_1,i_2,\ldots,i_k}\}$ having similar properties as $\{D_{i_1 i_2,\ldots,i_k}\}$. Write

$$D^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k) = D^{(n)}_{i_1,\ldots,i_k} \cap \phi_n^{-1} D_{j_1,\ldots,j_k}, \tag{3}$$

and

$$p^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k) = \mu_n(D^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)). \tag{4}$$

It is obvious that $d(D^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)) < 2^{-k}$.

Let $\Omega = [0,1)$, F be the σ-field of all Borel sets in Ω, and P be the Lebesgue measure restricted on F.

Split Ω into partitions $\{I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)\}$ with the following properties:

1) Each $I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$ is an interval closed from left and open from right, and has length $p^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$.

2) For each n and each k, $\{I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)\}$ is a partition of $[0,1)$.

(5)

3) $I^{(n)}_{i_1,\ldots,i_{k-1}}(j_1,\ldots,j_{k-1}) = \bigcup_{i_k=1}^{\infty} \bigcup_{j_k=1}^{\infty} I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$.

4) If $i_k < i'_k$, for any $i_1,\ldots,i_{k-1}$, $j_1,\ldots,j_k$, $j'_1,\ldots,j'_k$, $I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$ is located on the left of $I^{(n)}_{i_1,\ldots,i_{k-1},i'_k}(j'_1,\ldots,j'_k)$.

5) If $j_t < j'_t$, $t \leq k$, then for any $i_1,\ldots,i_k$, $j_1,\ldots,j_{t-1}$, $j_{t+1},\ldots,j_k$, $j'_{t+1},\ldots,j'_k$, $I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$ is located on the left of $I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_{t-1},j'_t,\ldots,j'_k)$.

We take a point $x^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$ artibrarily from $D^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$ if it is not empty and define

$$X^{(k)}_n(\omega) = x^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k), \text{ if } \omega \in I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k). \tag{6}$$

Evidently, for each n and k, X^k_n is measurable. Because

$$D^{(n)}_{i_1,\ldots,i_{k+1}}(j_1,\ldots,j_{k+1}) \subset D^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$$

and

$$d\left(D^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)\right) < 2^{-k},$$

we have

$$\rho_n\left(X^{(k)}_n(\omega), X^{(k+1)}_n(\omega)\right) < 2^{-k}. \tag{7}$$

Thus for each n, $\{X^{(k)}_n, k = 1,2,\ldots\}$ forms a Cauchy sequence and there exists a measurable function X_n such that

$$X^{(k)}_n \to X_n, \; k \to \infty, \; \forall\, \omega \in \Omega, \tag{8}$$

because S_n is complete. Therefore, we have defined an S_n-valued random element X_n for each n.

Next, we shall prove that μ_n is the distribution of X_n. Take any open set $A_n \subset S_n$, define $A_{m,n} = \{x^{(n)} \in A_n;\ \rho_n(x^{(n)}, \partial A_n) > \frac{1}{m}\}$, where m is a positive integer and ∂A_n is the boundary of A_n and $\rho_n(x^{(n)}, B) = \inf\{\rho_n(x^{(n)}, y^{(n)}),\ y^{(n)} \in B\}$. It is obvious that for each pair (n,m), $A_{m,n}$ is an open set contained in A_n, and $A_{m,n} \subset A_{m+1,n}$. Thus we have an expression of $A_{m,n}$ as follows

$$A_{nm} = \sum_{(i_1,j_1)\in N_{1,m}^{(n)}} D_{i_1}^{(n)}(j_1) + \sum_{(i_1,i_2;j_1,j_2)\in N_{2,m}^{(n)}} D_{i_1 i_2}^{(n)}(j_1,j_2) + \ldots \qquad (9)$$

where $N_{1,m}^{(n)}, N_{2,m}^{(n)}, \ldots$ are suitable index sets and all the right hand side terms are disjoint each other.

For each k with $2^{-k+1} < \frac{1}{2m^2}$, we have

$$\rho_n(X_n, X_n^{(k)}) < (\tfrac{1}{2})^{k-1} < \frac{1}{2m^2}. \qquad (10)$$

Hence

$$(X_n \in A_{m-1,n}) \subset [X_n^{(k)} \in A_{m,n}) \subset (X_n \in A_{m+1,n}).$$

Thus

$$P(X_n \in A_{m-1,n}) \leqq P(X_n^{(k)} \in A_{m,n}) \leqq P(X_n \in A_{m+1,n}). \qquad (11)$$

On the other hand, we have

$$P(X_n^{(k)} \in A_{m,n}) = \sum_{(i_1,j_1)\in N_{1,m}^{(n)}} P(X_n^{(k)} \in D_{i_1}^{(n)}(j_1))$$

$$+ \sum_{(i_1,i_2;j_1j_2)\in N_{2,m}^{(n)}} P(X_n^k \in D_{i_1 i_2}^{(n)}(j_1,j_2)) + \ldots$$

$$= \sum_{(i_1,j_1)\in N_{1,m}^{(n)}} |I_{i_1}^{(n)}(j_1)|$$

$$+ \sum_{(i_1,i_2;j_1,j_2)\in N_{2,m}^{(n)}} |I_{i_1 j_2}^{(n)}(j_1,j_2)| + \ldots$$

$$= \sum_{(i_1,j_1)\in N_{1,m}^{(n)}} \mu_n(D_{i_1}^{(n)}(j_1))$$

$$+ \sum_{(i_1,i_2;j_1,j_2)\in N_{2,m}^{(n)}} \mu_n(D_{i_1 i_2}^{(n)}(j_1,j_2)) + \ldots$$

$$= \mu_n(A_{m,n}). \tag{12}$$

From (11) and (12) it follows that

$$P(X_n \in A_{m-1,n}) \leq \mu_n(A_{m,n}) \leq P(X_n \in A_{m+1,n}) \leq P(X_n \in A_n). \tag{13}$$

If we let $m \to \infty$, we obviously have $A_{m-1,n} \uparrow A_n$, $A_{m,n} \uparrow A_n$. Hence from (13) we get

$$P(X_n \in A_n) = \mu_n(A_n) \tag{14}$$

Therefore, μ_n is the distribution of X_n, for each n.

Write

$$p_{i_1,\ldots,i_k} = \mu(D_{i_1,\ldots,i_k}),$$

and split Ω into partitions $\{I_{i_1,\ldots,i_k}\}$ such that

1) for each k, $\{I_{i_1,\ldots,i_k},\ i_1,\ldots,i_k = 1,2,\ldots\}$ is a partition of Ω,

2) each $I_{i_1,\ldots,i_k}$ is an interval closed from left and open from right, with its length $p_{i_1,\ldots,i_k}$,

3) $I_{i_1,\ldots,i_{k-1}} = \bigcup_{i_k=1}^{\infty} I_{i_1,\ldots,i_k}$,

4) if $i_k < i_k'$, then $I_{i_1,\ldots,i_k}$ is located on the left of $I_{i_1,\ldots,i_{k-1}i_k'}$.

We arbitrarily take a point $x_{i_1,\ldots,i_k}$ from $D_{i_1,\ldots,i_k}$ for each $(i_1,\ldots,i_k)$ and define

$$X^{(k)} = x_{i_1,\ldots,i_k}, \quad \text{if } \omega \in I_{i_1,\ldots,i_k}.$$

Similarly as before, we can prove that there exists an S-valued random element X such that

$$\rho(X^{(k)},X) < 2^{-k}, \tag{15}$$

and that μ is the distribution of X.

To complete the proof of the special case of Theorem 1, we need only to prove that

$$\phi_n(X_n) \to X, \quad \text{a.s.} \tag{16}$$

Write

$$I^{(n)}_{i_1,\ldots,i_k} = \bigcup_{j_1=1}^{\infty} \cdots \bigcup_{j_k=1}^{\infty} I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k).$$

According to the definition of $I^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k)$, $\{I^{(n)}_{i_1\ldots i_k}\}$ has analogous properties as $\{I_{i_1,\ldots,i_k}\}$, and their length satisfies

$$|I^{(n)}_{i_1\cdots i_k}| = \sum_{j_1=1}^{\infty} \cdots \sum_{j_k=1}^{\infty} \mu_n(D^{(n)}_{i_1,\ldots,i_k}(j_1,\ldots,j_k))$$

$$= \mu_n(\phi_n^{-1}(D_{i_1,\ldots,i_k})).$$

Since $\mu(\partial D_{i_1,\ldots,i_k}) = 0$ and $\mu_n\phi_n^{-1} \xrightarrow{W} \mu$, we have

$$|I^{(n)}_{i_1,\ldots,i_k}| \to \mu(D_{i_1,\ldots,i_k}), \quad \text{as} \quad n \to \infty. \tag{17}$$

If ω is a point of Ω and is not an endpoint of any interval $I_{i_1,\ldots,i_k}$, $k = 1,2,\ldots$, $i_1,\ldots,i_k = 1,2,\ldots$, then for each k there exists a k-multiple $(\alpha_1,\ldots,\alpha_k)$ such that ω is an inner point of $I_{\alpha_1,\ldots,\alpha_k}$. In view of the definition of $\{I_{i_1,\ldots,i_k}\}$, we know that the left and right endpoints of $I_{\alpha_1,\ldots,\alpha_k}$ are

$$a_k = \sum_{i_1=1}^{\alpha_1-1} \mu(D_{i_1}) + \sum_{i_2=1}^{\alpha_2-1} \mu(D_{\alpha_1 i_2}) + \ldots + \sum_{i_k=1}^{\alpha_k-1} \mu(D_{\alpha_1\cdots\alpha_{k-1}i_k})$$

and

$$b_k = a_k + \mu(D_{\alpha_1,\ldots,\alpha_k}).$$

Similarly, the two endpoints of $I^{(n)}_{\alpha_1,\ldots,\alpha_k}$ are

$$a_k^{(n)} = \sum_{i_1=1}^{\alpha_1-1} |I^{(n)}_{i_1}| + \sum_{i_2=1}^{\alpha_2-1} |I^{(n)}_{\alpha_1 i_2}| + \ldots + \sum_{i_k=1}^{\alpha_k-1} |I^{(n)}_{\alpha_1,\ldots,\alpha_{k-1}i_k}|$$

$$\leq \frac{\varepsilon}{2} + \sum_{j=1}^{\infty} \frac{(N_j+1)\varepsilon}{2^{j+1}(N_1,\ldots,N_j+1)} < \varepsilon.$$

Finally, we shall prove that

$$\mu\left(x \in S_1 \ , \ \phi_\varepsilon(x) \ \text{ is discontinuous at } \ x\right) = 0.$$

If $x \in \left(K_0^{k_0}\right)^0 \bigcap K_{\alpha_1,\ldots,\alpha_{k_0}-1}$, for some k_0 and $\alpha_1,\ldots,\alpha_{k_0}-1$, according to the definition of $\phi_\varepsilon(x)$

$$\phi_\varepsilon(z) = y_{\alpha_1,\ldots,\alpha_{k_0}-1} \ , \text{ for any } \ z \in K_0^{k_0} \bigcap K_{\alpha_1,\ldots,\alpha_{k_0}-1}.$$

Hence $\phi(x)$ is continuous at x. If $x \in \bigcup_{k=1}^{\infty} K_0^k$, then for any k, there exists a k-ple $(\alpha_1,\ldots,\alpha_k)$ such that $x \in K_{\alpha_1,\ldots,\alpha_k}$, $\alpha_1 \leq N_1,\ldots,\alpha_k \leq N_k$. Since $K_{\alpha_1,\ldots,\alpha_k}$ is an open set, we have that $\rho\left(x, \partial K_{\alpha_1,\ldots,\alpha_k}\right) = \rho_k > 0$. If $y \in S_1$ and $\rho(x,y) < \rho_k$, then $y \in K_{\alpha_1,\ldots,\alpha_k}$. Hence

$$\rho\left(\phi_k(x) \ , \ \phi_\varepsilon(x)\right) < \frac{1}{2^{k-1}}$$

$$\rho\left(\phi_k(y) \ , \ \phi_\varepsilon(y)\right) < \frac{1}{2^{k-1}}$$

and

$$\phi_k(x) = \phi_k(y).$$

Therefore,

$$\rho\left(\phi_\varepsilon(x) \ , \ \phi_\varepsilon(y)\right) < 1/2^{k-2}.$$

$$< \frac{\varepsilon}{2} + \sum_{j=1}^{k-1} \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_j=1}^{N_j} \frac{(i_j+1)\varepsilon}{2^{j+1}(N_1,\ldots,N_j+1)^2}$$

$$< \frac{\varepsilon}{2} + \sum_{j=1}^{k-1} \frac{(N_j+1)\varepsilon}{2^{j+1}(N_1,\ldots,N_j+1)}$$

Thus

$$\mu\left(\bigcup_{k=1}^{\infty} K_0^k\right) < \frac{\varepsilon}{2} + \sum_{j=1}^{\infty} \frac{(N_j+1)\varepsilon}{2^{j+1}(N_1,\ldots,N_j+1)} \quad . \tag{31}$$

On the other hand, we have

$$\mu\left(\rho\left(\phi_k(x)\ ,\ \phi(x)\right) > \frac{1}{2^{k-1}}\ ,\ x \in \bigcup_{i_1=1}^{N_1}, \ldots, \bigcup_{i_k=1}^{N_{k_j}} K_{i_1,\ldots,i_k}\right)$$

$$\leq \sum_{i=1}^{N_1}, \ldots, \sum_{i_k=1}^{N_k} \mu(K_{i_1,\ldots,i_k} \Delta E_{i_1,\ldots,i_k})$$

$$\leq \sum_{i=1}^{N_1}, \ldots, \sum_{i_k=1}^{N_k} \sum_{j=1}^{i_k} \mu(G_{i_1,\ldots,i_{k-1}\,j} \Delta E_{i_1,\ldots,i_{k-1}j})$$

$$\leq N_1,\ldots,N_k \cdot N_k \left(\varepsilon/2^{k+1}(N_1,\ldots,N_k+1)^2\right) \leq \varepsilon/2^{k+1} \tag{32}$$

By (30) (31) (32), we obtain

$$\mu\left(\phi_\varepsilon(x) = \phi(x)\right) = \lim_{k\to\infty} \mu\left(\rho\left(\phi_\varepsilon(x)\ ,\ \phi(x)\right) > \frac{1}{2^{k-2}}\right)$$

$$= \frac{\varepsilon}{2} + \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_{k-1}=1}^{N_{k-1}} \mu(E_{i_1,\ldots,i_{k-1}} \setminus K_{i_1,\ldots,i_{k-1}})$$

$$+ \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_{k-2}=1}^{N_{k-2}} \mu(\phi^{-1}D_{i_1,\ldots,i_{k-2}} \setminus K_{i_1,\ldots,i_{k-2}}) \leq \ldots$$

$$\leq \frac{\varepsilon}{2} + \sum_{j=1}^{k-1} \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_j=1}^{N_j} \mu(E_{i_1,\ldots,i_j} \setminus K_{i_1,\ldots,i_j})$$

$$\leq \frac{\varepsilon}{2} + \sum_{j=1}^{k-1} \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_j=1}^{N_j} \left[\mu(G_{i_1,\ldots,i_j} \Delta E_{i_1,\ldots,i_j}) \right.$$

$$\left. + \mu(G_{i_1,\ldots,i_j} \setminus K_{i_1,\ldots,i_j}) \right]$$

$$\leq \frac{\varepsilon}{2} + \sum_{j=1}^{k-1} \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_j=1}^{N_j} \left[\frac{\varepsilon}{2^{j+1}(N_1,\ldots,N_j+1)^2} \right.$$

$$\left. + \mu(G_{i_1,\ldots,i_j} \bigcap \bigcup_{\ell=1}^{i_j-1} G_{i_1,\ldots,i_{j-1}\ell}) \right]$$

$$\leq \frac{\varepsilon}{2} + \sum_{j=1}^{k-1} \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_j=1}^{N_j} \left[\frac{\varepsilon}{2^{j+1}(N_1,\ldots,N_j+1)^2} \right.$$

$$\left. + \sum_{\ell=1}^{i_j} \mu(E_{i_1,\ldots,i_{j-1}\ell} \Delta G_{i_1,\ldots,i_{j-1}\ell}) \right]$$

We have

$$\mu\Big(\bigcup_{j=1}^{k} K_0^j\Big) \le \mu\Big(S_1 \setminus \bigcup_{i_1=1}^{N_1}, \ldots, \bigcup_{i_k=1}^{N_k} K_{i_1,\ldots,i_k}\Big)$$

$$= \mu\Big(S_1 \setminus \bigcup_{i_1=1}^{N_1}, \ldots, \bigcup_{i_k=1}^{N_k} G_{i_1,\ldots,i_k}\Big)$$

$$\le \mu\Big(S_1 \setminus \bigcup_{i_1=1}^{N_1}, \ldots, \bigcup_{i_k=1}^{N_k} E_{i_1,\ldots,i_k}\Big) + \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_k=1}^{N_k} \mu(G_{i_1,\ldots,i_k} \Delta E_{i_1,\ldots,i_k})$$

$$\le \mu\Big(S_1 \setminus \bigcup_{i_1=1}^{N_1}, \ldots, \bigcup_{i_k=1}^{N_k} \phi^{-1} D_{i_1,\ldots,i_k}\Big)$$

$$+ \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_k=1}^{N_k} \mu(\phi^{-1} D_{i_1,\ldots,i_k} \setminus K_{i_1,\ldots,i_{k-1}})$$

$$+ N_1,\ldots,N_k \ \varepsilon / 2^{k+1}(N_1,\ldots,N_k + 1)^2$$

$$< \frac{\varepsilon}{2} - \frac{\varepsilon}{2^{k+1}} + \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_{k-1}=1}^{N_{k-1}} \mu(\phi^{-1} D_{i_1,\ldots,i_{k-1}} \setminus K_{i_1,\ldots,i_{k-1}})$$

$$+ \frac{\varepsilon}{2^{k+1}(N_1,\ldots,N_k + 1)}$$

$$< \frac{\varepsilon}{2} + \sum_{i=1}^{N_1}, \ldots, \sum_{i_{k-1}=1}^{N_{k-1}} \mu(E_{i_1,\ldots,i_{k-1}} \setminus K_{i_1,\ldots,i_{k-1}})$$

$$+ \sum_{i_1=1}^{N_1}, \ldots, \sum_{i_{k-1}=1}^{N_{k-1}} \mu(\phi^{-1} D_{i_1,\ldots,i_{k-1}} \setminus K_{i_1,\ldots,i_{k-2}})$$

By the definition of ϕ_k and ϕ_{k+1} we get that

$$\phi_k(x) = y_{\alpha_1,\ldots,\alpha_k} \in D_{\alpha_1,\ldots,\alpha_k}$$

$$\phi_{k+1}(x) = y_{\alpha_1,\ldots,\alpha_k\,\beta_{k+1}} \in D_{\alpha_1,\ldots,\alpha_k\,\beta_{k+1}} \subset D_{\alpha_1,\ldots,\alpha_k}.$$

Therefore

$$\rho\Big(\phi_k(x),\ \phi_{k+1}(x)\Big) < 2^{-k}$$

Thus there must be a limit point, denoted by $\phi_\varepsilon(x)$, of the sequence $\phi_k(x)$.

Combining this and (26) we obtain that $\lim_{k\to\infty} \phi_k(x) = \phi_\varepsilon(x)$ pointwise. (27)

Now we shall prove that

$$\mu\Big(\phi_\varepsilon(x) \neq \phi(x)\Big) < \varepsilon. \tag{28}$$

Note that if $x \in \bigcup_{k=1}^{\infty} K_0^k$, we always have

$$\rho\Big(\phi_\varepsilon(x)\ ,\ \phi_k(x)\Big) < 1/2^{k-1}, \text{ for any } k \geq 1. \tag{29}$$

Therefore, for any k

$$\mu\Big(\rho\Big(\phi_\varepsilon(x)\ ,\ \phi(x)\Big) > 1/2^{k-2}\Big) \leq \mu\Big(\bigcup_{k=1}^{\infty} K_0^k\Big) +$$

$$+ \mu\Big(\rho\Big(\phi_k(x)\ ,\ \phi(x)\Big) > \frac{1}{2^{k-1}}\ ,\ x \in \bigcup_{i_1=1}^{N_1},\ldots,\bigcup_{i_k=1}^{N_k} K_{i_1,\ldots,i_k}\Big). \tag{30}$$

5) $K_{i_1,\ldots,i_{k-1}1} = G_{i_1,\ldots,i_{k-1}1}$, $K_{i_1,\ldots,i_{k-1}2} = (G_{i_1,\ldots,i_{k-1}2} \setminus G_{i_1,\ldots,i_{k-1}1})$

$,\ldots,$ $K_{i_1,\ldots,i_{k-1}N_k} = (G_{i_1,\ldots,i_{k-1}N_k} \setminus \bigcup_{i_k=1}^{N_k-1} G_{i_1,\ldots,i_k})^0$.

6) $K_0^k = \bigcup_{i_1=1}^{N_1} \cdots \bigcup_{i_{k-1}=1}^{N_{k-1}} (K_{i_1,\ldots,i_{k-1}} \setminus \bigcup_{i_k=1}^{N_k} K_{i_1,\ldots,i_k})$

7) $$\phi_k(x) = \begin{cases} y_{i_1,\ldots,i_k} & \text{if} \quad x \in K_{i_1,\ldots,i_k} \quad 1 \le i, \le N_1,\ldots,1 \le i_k \le N_k \\ \phi_{k-1}(x) & \text{if} \quad x \in K_0^1 \cup \ldots \cup K_0^k. \end{cases}$$

If $x \in K_0^{k_0}$ for some k_0, then for any $k > k_0$

$$\phi_k(x) = \phi_{k_0}(x) \tag{26}$$

because $K_0^k \bigcup_{i_1=1}^{N_1},\ldots,\bigcup_{i_k=1}^{N_k} K_{i_1,\ldots,i_k} \subset S_1 \setminus K_0^{(k-1)} \cup,\ldots,\cup K_0^1$. If $x \notin \bigcup_{k=1}^{\infty} K_0^k$, then for any k, $x \in \bigcup_{i_1=1}^{N_1} \cdots \bigcup_{i_k=1}^{N_k} K_{i_1,\ldots,i_k}$. Suppose that $x \in K_{\alpha_1,\ldots,\alpha_k}$, and $x \in K_{\beta_1,\ldots,\beta_{k+1}}$. Since $K_{\beta_1,\ldots,\beta_{k+1}} \subset G_{\beta_1,\ldots,\beta_{k+1}} \subset K_{\beta_1,\ldots,\beta_k}$ and $K_{i_1,\ldots,i_k}$'s are disjoint, it follows that $\beta_1 = \alpha_1,\ldots, \beta_k = \alpha_k$.

Write

$$K_{i_1 1} = G_{i_1 1},\ K_{i_1 2} = (G_{i_1 2} \setminus G_{i_1 1})^0, \ldots,\ K_{i_1 N_2} = (G_{i_1 N_2} \setminus \bigcup_{i_2=1}^{N_2-1} G_{i_1 i_2})^0,$$

and

$$K_0^2 = \bigcup_{i_1=1}^{N_2} (K_{i_1} \setminus \bigcup_{i_2=1}^{N_2} K_{i_1 i_2}).$$

Define

$$\phi_2(x) = \begin{cases} \overline{y_{i_1 i_2}} & \text{if } x \in K_{i_1 i_2},\ 1 \le 1_1 \le N_1,\ 1 \le i_2 \le N_2, \\ \phi_1(x) & \text{if } x \in K_0^1 \cup K_0^2. \end{cases}$$

Then let $E_{i_1 i_2 i_3} = K_{i_1 i_2} \cap \phi^{-1} D_{i_1 i_2 i_3}$, $1 \le i_1 \le N_1$, $1 \le i_2 \le N_2$, $i \le i_3 \le N_3$. Similarly define $G_{i_1 i_2 i_3}$, $K_{i_1 i_2 i_3}$, K_0^3 and $\phi_3(x)$. By induction we can define $E_{i_1,\ldots,i_k}$, $G_{i_1 i_2,\ldots,i_k}$, $K_{i_1,\ldots i_k}$, K_0^k, $\phi_k(x)$ satisfying the following relations:

1) $E_{i_1,\ldots,i_k} = K_{i_1,\ldots,i_{k-1}} \cap \phi^{-1} D_{i_1,\ldots,i_k}$, $1 \le i_1 \le N_1,\ldots,1 \le i_k \le N_k$.

2) $G_{i_1,\ldots,i_k} \subset K_{i_1,\ldots,i_{k-1}}$, $G_{i_1,\ldots,i_k}$'s are open sets.

3) $\mu(G_{i_1,\ldots i_k} \Delta E_{i_1,\ldots,i_k}) < \varepsilon/2^{k+1}(N_1,\ldots,N_k+1)^2$

4) $\mu(\partial G_{i_1,\ldots,i_k}) = 0$

Let $E_{i_1} = \phi^{-1}D_{i_1}$. For each i_1, there is an open set G_{i_1} such that

$$\mu(\partial G_{i_1}) = 0$$

and

$$\mu(G_{i_1} \Delta E_{i_1}) < \varepsilon/4(N_1+1)^2.$$

Write

$$K_1 = G_1, \quad K_2 = (G_2 \setminus G_1)^0, \ldots, \quad K_{N_1} = (G_1 \setminus \bigcup_{i_1=1}^{N_1-1} G_{i_1})^0$$

and $K_0^1 = S_1 \setminus \bigcup_{i=1}^{N_1} K_i$, where A^0 denotes the interior of the set A. Define

$$\phi_1(x) = \begin{cases} y_{i_1} & \text{if } x \in K_{i_1}, \ i_1 = 1,\ldots,N_1, \\ y^0 & \text{if } x \in K_0^1. \end{cases}$$

Secondly, let $E_{i_1 i_2} = K_{i_1} \cap \phi^{-1}(D_{i_1 i_2})$. Then there exist open sets $G_{i_1 i_2}$, $i_1 \le N_{i_1}$, $i_2 \le N_2$, such that

1) $G_{i_1 i_2} \subset K_{i_1}$, $\quad i_1 = 1,2,\ldots,N_1$, $i_2 = 1,2,\ldots,N_2$,

2) $\mu(G_{i_1 i_2} \Delta E_{i_1 i_2}) < \varepsilon/8(N_1 N_2+1)^2$,

3) $\mu(\partial G_{i_1 i_2}) = 0$.

$$\phi_n(X_n) \to X, \quad \text{a.s.}, \ n \to \infty. \tag{25}$$

As before, we can make a slight modification on X_n and X so that (25) holds pointwise. Theorem 1 is proved.

Now we turn to prove Theorem 2. Suppose that S_1 and S_2 are two complete separable metric spaces, μ is a finite measure defined on S_1, ϕ: $S_1 \to S_2$ is a measurable mapping.

Using the same approach, we split S_2 into a sequence of partitions $\{D_{i_1,\ldots,i_k},\ i_1,\ldots,i_k = 1,2,\ldots\}$, $k = 1,2,\ldots$ such that

$$D_{i_1,\ldots,i_{k-1}} = \bigcup_{i_k=1}^{\infty} D_{i_1,\ldots,i_k}, \quad k = 2,3,\ldots$$

$$S_2 = \bigcup_{i_1=1}^{\infty} D_{i_1},$$

and $d(D_{i_1,\ldots,i_k}) < 1/2^{-k}$. For any fixed $\varepsilon > 0$, we can select a sequence of positive integers $N_1, N_2, \ldots$, such that

$$\sum_{i_1=1}^{N_1} \cdots \sum_{i_k=1}^{N_k} \mu(\phi^{-1} D_{i_1,\ldots,i_k}) > \mu(S_1) - \left(\frac{\varepsilon}{4} + \frac{\varepsilon}{8} + \ldots + \frac{\varepsilon}{2^{k+1}}\right) = \mu(S_1) - \frac{\varepsilon}{2} + \frac{\varepsilon}{2^{k+1}},$$

for any $k = 1,2,\ldots$. Without any loss of generality, we can assume that each $D_{i_1,\ldots i_k}$ is nonempty, $i_1 = 1,\ldots,N_1,\ldots,i_k = 1,\ldots,N_k$. Arbitrarily take $y_{i_1,\ldots,i_k} \in D_{i_1,\ldots,i_k}$, $i_1 \le N_1,\ldots i_k \le N_k$, and $y^0 \in S_2$.

On the other hand, it is obvious that

$$\tilde{\phi}_n^{-1}B \,\Delta\, \phi_n^{-1}\, B \subset D_n$$

where $A \,\Delta\, B$ denotes $(A\setminus B)\cup(B\setminus A)$. Thus we have

$$|\mu_n\phi_n^{-1}\, B \;-\; \mu_n\tilde{\phi}_n^{-1}\, B\,| \le \mu_n(\tilde{\phi}_n^{-1}B \,\Delta\, \phi_n^{-1}\, B\,) \le \mu_n(D_n) \le \frac{1}{2^n} \to 0,$$

$$\text{as} \quad n \to \infty.$$

Therefore (24) follows from the above estimate and the fact that $\mu_n\phi_n^{-1} \xrightarrow{W} \mu$

According to the special case that we just proved, we can find a probability space (Ω, F, P) on which there is a sequence of random elements X_n and X such that

1) μ_n is the distribution of X_n, μ is the distribution of X,
2) $\tilde{\phi}_n(X_n) \to X$ pointwise.

Since

$$\sum_{n=1}^{\infty} P(\omega:\ \tilde{\phi}_n(X_n(\omega)) \neq \phi_n(X(\omega)))$$

$$= \sum_{n=1}^{\infty} \mu_n(x \in S_n\ \ \tilde{\phi}_n(x) \neq \phi_n(x))$$

$$\le \sum_{n=1}^{\infty} \frac{1}{2^n} = 1 < \infty,$$

by Borel-Cantelli lemma we know that

Example: Let $S_1 = [0,1]$ with Euclidean norm and Lebesgue measure, and let $S_2 = \{0,1\}$ with $\rho(0,1) = 1$. Define

$$\phi = I_{[0,\frac{1}{2}]}(x), \quad x \in S_1$$

where I_A denotes the indicator of the set A. For any $\varepsilon < \frac{1}{2}$, we cannot find a continuous mapping $\phi_\varepsilon: S_1 \to S_2$ satisfying (22).

A little more complex example yields from the above example with the measure replaced by μ:

$$\mu(B) = \frac{1}{2} L(B) + \sum_{r_n \in B} \frac{1}{2^{n+1}}, \qquad B \in \mathcal{B}([0.1]),$$

where $L(B)$ is the Lebesque measure of the set B and $Q = \{r_n, n = 1,2,\ldots\}$ is the set of all rational numbers in $[0,1]$.

Before we prove Theorem 2, we first use Theorem 2 to complete the proof of Theorem 1. For each n, according to Theorem 2, there exists a measurable mapping $\tilde{\phi}_n$ such that

1) $\mu_n(\phi_n \neq \tilde{\phi}_n) < 1/2^n$,

2) $\mu_n(x \in S_n; \tilde{\phi}_n \text{ is discontinuous at } x) = 0$.

We shall first prove that

$$\mu_n \tilde{\phi}_n^{-1} \xrightarrow{W} \mu. \tag{24}$$

Let B be a Borel subset of S and $B_n = \phi_n^{-1}B$, $\tilde{B}_n = \tilde{\phi}_n^{-1}B$. Denote by $D_n = \{x \in S_n, \phi_n(x) \neq \tilde{\phi}_n(x)\}$. By the definition of $\tilde{\phi}_n$, we have $\mu_n(D_n) < \frac{1}{2^n}$.

Since $P(N) = 0$, $\tilde{X}_n$ and X_n (correspondingly $\tilde{X}$ and X) have the same distribution. Thus Theorem 1 holds when ϕ_n are all continuous.

Note that the continuity of ϕ_n is only used in deriving that

$$\rho(\phi_n(X_n),\phi_n(X_n^{(m)})) \to 0, \text{ as } m \to \infty,$$

(see (20) and (21)). We can relax the continuity restriction as the following

$$\mu_n\{x \in S_n;\ \phi_n \text{ is discontinuous at } x\} = 0,$$

for each n. Therefore, to complete the proof of Theorem 1, we only need the following generalized Lusin's Theorem.

THEOREM 2: Let S_1 and S_2 be two complete separable metric spaces, μ be a finite measure defined S_1 and let ϕ be a measurable mapping from S_1 into S_2. Then for any $\varepsilon > 0$, there exists a measurable mapping ϕ_ε: $S_1 \to S_2$, satisfying

1) $\mu(\phi \neq \phi_\varepsilon) < \varepsilon$, (22)

2) $\mu(x \in S_1,\ \phi_\varepsilon \text{ is discontinuous at } x) = 0$. (23)

Remark: The main difference between Theorem 2 and the ordinary Lusin's Theorem is the condition (23). But in the general case, we cannot require that ϕ_ε is continuous. This can be seen from the following example:

From this and $\phi_n(X_n^{(k)}(\omega)) \in D_{\alpha_1 \ldots \alpha_k}$, we get

$$\neg(\phi_n(X_n^{(m)}(\omega)),\ \phi_n(X_n^{(k)}(\omega)) < 2^{-k}. \tag{20}$$

From (19) and (20), we get

$$\rho(\phi_n(X_n(\omega)),X(\omega)) \leq 3\cdot 2^{-k} + \rho(\phi_n(X_n(w)),\phi_n(X_n^{(m)}(\omega))). \tag{21}$$

Since ϕ_n is continuous and $X_n^{(m)}(\omega) \to X_n(\omega)$, $m \to \infty$, it follows that

$$\rho(\phi_n(X_n(\omega)),X(\omega)) \leq 3\cdot 2^{-k}.$$

This proves (16) because the set of all endpoints of $I_{i_1,\ldots,i_k}$, $k = 1,2,\ldots$, $i,\ldots,i_k = 1,2,\ldots$, is countable, hence its Lebesgue measure is zero.

Let $N \subset \Omega$ be the set on which $\phi_n(X_n(\omega))$ do not converge to $X(\omega)$ and let $\phi_n(X_n(\omega_0)) \to X(\omega_0)$. Define a new sequence of random elements as follows:

$$\tilde{X}_n(\omega) = \begin{cases} X_n(\omega) & \text{if} \quad \omega \in \Omega \setminus N \\ X_n(\omega_0) & \text{if} \quad \omega \in N \end{cases}$$

and

$$\tilde{X}(\omega) = \begin{cases} X(\omega) & \text{if} \quad \omega \in \Omega \setminus N \\ X(\omega_0) & \text{if} \quad \omega \in N \end{cases}$$

Then we have

$$\phi_n(\tilde{X}_n(\omega)) \to \tilde{X}(\omega) \quad \text{pointwise.}$$

and

$$b_k^{(n)} = a_k^{(n)} + |I^{(n)}_{\alpha_1 \cdots \alpha_k}|.$$

From (17) we have that

$$a_k^{(n)} \to a_k \qquad (n \to \infty)$$

and

$$b_k^{(n)} \to b_k \qquad (n \to \infty)$$

Therefore, when n is large enough, $\omega \in I^{(n)}_{\alpha_1,\ldots,\alpha_k}$. Hence

$$\phi_n(X_n^{(k)}(\omega)) \in D_{\alpha_1 \cdots \alpha_k}.$$

Note that $X^{(k)}(\omega) \in D_{\alpha_1,\ldots,\alpha_k}$, we get

$$\rho(\phi_n(X_n^{(k)}(\omega)),\ X^{(k)}(\omega)) < 2^{-k}. \tag{18}$$

From (15) and (18) it follows that

$$\rho(\phi_n(X_n^{(k)}(\omega)),\ X(\omega)) < 2^{-k+1}. \tag{19}$$

For fixed n and k, and for any $m > k$, there exist $\alpha_{k+1},\alpha_{k+2},\ldots,\alpha_m$

Such that $\omega \in I^{(n)}_{\alpha_1,\ldots,\alpha_k,\alpha_{k+1},\ldots,\alpha_m}$. Thus we have

$$X_n^{(m)}(\omega) \in \phi_n^{-1}(D_{\alpha_1,\ldots,\alpha_m}) \subset \phi_n^{-1}(D_{\alpha_1,\ldots,\alpha_k})$$

or

$$\phi_n(X_n^{(m)}(\omega)) \in D_{\alpha_1,\ldots,\alpha_k}.$$

Since k is arbitrary, we have proved that $\phi_\varepsilon(x)$ is continuous at x. Thus

$$\mu(x \in S\ ,\ \phi_\varepsilon(x) \text{ is discontinuous at } x) \leq \sum_{k=1}^{\infty} \sum_{i_1=1}^{N_1} \cdots \sum_{i_k=1}^{N_k} \mu(\partial K_{i_1,\ldots,i_k}) = 0.$$

This completes the proof of Theorem 2.

3. A SIMPLE PROOF OF THEOREM 1 FOR THE FINITE DIMENSION CASE.

3.1 ONE DIMENSION CASE

Suppose that F_n and F are one-dimensional distributions satisfying that $F_n \xrightarrow{W} F$ as $n \to \infty$. Let $\Omega = (0,1)$. $\mathcal{F} = \mathcal{B}(0,1)$ and P be the Lebesque measure restricted on Ω. Define

$$X_n(\omega) = \text{Sup}\{x: F_n(x) < \omega\}\ ,\ \omega \in \Omega = (0,1),$$

and

$$X(\omega) = \text{Sup}\{x: F(x) < \omega\}.$$

According to this definition, it is evident that $X_n(\omega) \leq x$, $\omega \in \Omega$, $x \in R'$ is equivalent to the fact that $F_n(x) \geq \omega$. This ensures that F_n is the distribution of X_n. Similarly, F is the distribution of X.

For any $\omega \in (0,1)$, take arbitrarily $x_0 < X(\omega)$ and x_0 is a continuous point of $F(x)$. Then $F(x_0) < \omega$. Since x_0 is a continuous point of $F(x)$ and $F_n \xrightarrow{\omega} F$, we have $F_n(x_0) \to F(x_0)$. Thus when n is large enough we have $F_n(x_0) < \omega$, so that $X_n(\omega) \geq x_0$. Hence $X(\omega) \leq \underline{\lim}_{n\to\infty} X_n(\omega)$ for any $\omega \in (0,1)$.

Let $\omega \in (0,1)$ be such that for any $\varepsilon > 0$, $F\ X(\omega) + \varepsilon\ > \omega$. Take $\varepsilon > 0$ such that $X(\omega) + \varepsilon$ is a continuous point of $F(x)$. Since $F_n\left(X(\omega) + \varepsilon\right) \to F\left(X(\omega) + \varepsilon\right) > \omega$, when n is large enough we have $F_n\left(X(\omega) + \varepsilon\right) > \omega$. Thus, according to the definition of $X_n(\omega)$, $X_n(\omega) \leq X(\omega) + \varepsilon$. This shows that $\overline{\lim}_{n\to\infty} X_n(\omega) \leq X(\omega)$. Hence, $\lim_{n\to\infty} X_n(\omega) = X(\omega)$. If for some $\omega \in (0,1)$, there exists a constant $\varepsilon_0 > 0$ such that $\omega = F\left(X(\omega) + \varepsilon_0\right)$, then for any $0 < \varepsilon < \varepsilon_0$, $F\left(X(\omega) + \varepsilon_0\right) \geq F\left(X(\omega) + \varepsilon\right) \geq \omega = F\left(X(\omega) + \varepsilon_0\right)$ Hence $F\left(X(\omega) + \varepsilon\right) = F\left(X(\omega) + \varepsilon_0\right) = \omega$. Thus there exists a rational number $\gamma = \gamma(\omega) \in \left(X(\omega)\ ,\ X(\omega) + \varepsilon_0\right)$, corresponding to ω. If there are two points $\omega_1 < \omega_2$ which correspond to γ_1, γ_2 respectively, we shall prove that $\gamma_1 < \gamma_2$. In fact, if $\omega_i = F\left(X(\omega_i) + \varepsilon_i\right)$, $\varepsilon_i > 0$, $i = 1, 2$, then $X(\omega_1) + \varepsilon_1 \leq X(\omega_2)$. Otherwise, $X_1(\omega_1) \leq X(\omega_2) < X(\omega_1) + \varepsilon_1 \leq X(\omega_2) + \varepsilon_2$ would imply that $\omega_1 = F\left(X(\omega_1) + \varepsilon_1\right) = F\left(X(\omega_2) + \varepsilon_2\right) = \omega_2$, contradicting to the assumption that $\omega_1 < \omega_2$. Thus there are at most countably many ω such that $X_n(\omega) \not\to X(\omega)$. Hence

$$X_n(\omega) \longrightarrow X(\omega) \quad \text{a.s.}, \quad n \to \infty.$$

As before, we can change the definition of $X_n(\omega)$ and of $X(\omega)$ at those ω's at which $X_n(\omega) \not\to X(\omega)$, so that

$$X_n(\omega) \longrightarrow X(\omega) \quad \text{pointwise}, \quad n \to \infty.$$

3.2 TWO DIMENSION CASE

Let $F^{(n)}(\cdot\ ,\ \cdot)$ and $F(\cdot\ ,\ \cdot)$ be two-dimensional distributions such that $F^{(n)} \xrightarrow{w} F$, $n \to \infty$. Let $F_X^{(n)}(\cdot)$ and $F_y^{(n)}(\cdot|x)$ denote the marginal distribution of the first component and the conditional distribution of the second component when given the first component to be x, corresponding to $F^{(n)}$. Define random vectors $(X_n\ ,\ Y_n)$ on $\Omega = (0,1) \times (0,1)$ as follows:

$$\begin{cases} X_n(\omega_1\ ,\ \omega_2) \overset{\Delta}{=} X_n(\omega_1) \overset{\Delta}{=} \operatorname{Sup}\left\{x:\ F_X^{(n)}(x) < \omega_1\right\} & , \text{ if } \omega_1 \in (0,1) \\ Y_n(\omega_1\ ,\ \omega_2) = \operatorname{Sup}\left\{y:\ F_Y^{(n)}\left(y|X_n(\omega_1)\right) < \omega_2\right\} & , \text{ if } \omega_i \in (0,1),\ i = 1,2. \end{cases}$$

We can similarly define random vector $\{X,\ Y\}$. As before we can show that $\{X_n(\omega_1) \le X\ ,\ Y_n(\omega_1,\omega_2) \le y\}$ is equivalent to $\{F_X^{(n)}(x) \ge \omega_1\ ,\ F_y^{(n)}\left(y|X_n(\omega_1)\right) \ge \omega_2\}$, and that

$$P(X_n \le x\ ,\ Y_n \le y) = \int_0^{F_X^{(n)}(x)} F_y^{(n)}\left(y|X_n(\omega_1)\right)\ d\ \omega_1$$

$$= \int_{-\infty}^{x} F_y^{(n)}(y|t)\ F_X^{(n)}(dt) = F_n(x\ ,\ y).$$

Similarly, $F(\cdot\ ,\ \cdot)$ is the distribution of $(X\ ,\ Y)$. Using the conclusion in 3.1, we have

$$X_n(\omega_1) \longrightarrow X(\omega_1) \quad \text{a.s.}$$

with respect to the one-dimensional Lebesque measure restricted on $(0,1)$.

By Fubini's Theorem, we know that

$$X_n(\omega_1,\omega_2) = X_n(\omega_1) \longrightarrow X(\omega_1) = X(\omega_1,\omega_2)$$

with respect to the two-dimensional Lebesque measure restricted on $(0,1) \times (0,1)$.

Again using the conclusion about one dimension case, for any fixed $\omega_1 \in (0,1)$, we get that

$$Y_n(\omega_1,\omega_2) \longrightarrow Y(\omega_1,\omega_2) \quad \text{a.s.}$$

with respect to the one-dimensional Lebesque measure restricted on $(0,1)$.

Again using Fubini's Theorem, we obtain

$$Y_n(\omega_1,\omega_2) \longrightarrow Y(\omega_1,\omega_2) \quad \text{a.s.}$$

with respect to the two-dimensional Lebesque measure restricted on $(0,1) \times (0,1)$.

For d-dimension case, the proof is the same as in two-dimension case.

4. APPLICATIONS OF THEOREM 1

4.1 HELLEY-BRAY THEOREM ([2] and [4]).

If $F_n \xrightarrow{w} F$ and $g(x)$ is a continuous bounded function, then

$$\int g(x)F_n(dx) \longrightarrow \int g(x)F(dx)$$

PROOF. Construct $X_n \sim F_n$, $X \sim F$ and $X_n \longrightarrow X$, according to Theorem 1. Then by the dominated convergence theorem we have

$$\int g(x)F_n(dx) = Eg(X_n) \longrightarrow Eg(X) = \int g(x)F(dx)$$

4.2 (See [4])

If $F_n \xrightarrow{w} F$, then $f_n(t) \longrightarrow f(t)$ uniformly on any bounded interval, where f_n and f are the characteristic functions of F_n and F, respectively.

PROOF.

Let $T > 0$ be any fixed number. Then

$$\begin{aligned} |f_n(t) - f(t)| &= |E(e^{itX_n} - e^{itX})| \\ &\leq E|e^{it(X_n - X)} - 1| \\ &\leq 2P(|X_n - X| \geq \varepsilon/T) + \varepsilon \longrightarrow \varepsilon \ , \ \forall\, |t| \leq T. \end{aligned}$$

Hence $|f_n(t) - f(t)| \longrightarrow 0$ uniformly on $[-T, T]$.

4.3 (See [4]).

If $F_n \xrightarrow{w} F$ and $r > 0$, then

$$\int |X|^r F(dx) \leq \varliminf_{n\to\infty} \int |X|^r F_n(dx).$$

PROOF.

Let $X_n \sim F_n$, $X \sim F$ and $X_n \to X$. Then what to be proved is equivalent to

$$E|X|^r \leq \varliminf_{n\to\infty} E|X_n|^r.$$

The latter is just a special case of Fatou Lemma.

4.4

If $\{X_n\}$ converges in distribution to F, Y_n to E_a, the degenerate distribution concentrating its mass at a, and Z_n converges to E_b, $b > 0$, then $\{(X_n + Y_n)Z_n\}$ converges in distribution to $G(x) = F(\frac{x}{b} - a)$ (See [4], Th.4.4.6 and the corollary after it).

PROOF.

Since $\{(X_n , Y_n , Z_n)\}$ converges in distribution to $F(x)E_a(y)E_b(z)$. By Theorem 1, we can construct $(Z_n , Y_n , Z_n) \longrightarrow (Z , a , b) \sim F(x)E_a(y)E_b(z)$. Thus $(Z_n + Y_n)Z_n \longrightarrow (z+a)b \sim F(\frac{x}{b} - a)$. Q.E.D.

Though the original proofs of the above four results are not very complicated, the proofs given here are relatively easier. In the following examples, the proofs will be involved with much difficulty if you do not use Theorem 1.

4.5 (See [1], [7] and [8]).

Suppose that $\{W_{ij}^m,\ 1 \le i \le p,\ 1 \le j \le k-1\} \xrightarrow{\text{in d.}} \{W_{ij},\ 1 \le 1 \le P,\ 1 \le j \le k-1\}$ and that $\{U_{ij}^{(m)},\ 1 \le i \le j \le P\} \xrightarrow{\text{in d.}} \{U_{ij},\ 1 \le 1 \le j \le P\}$. Consider the detrimental equation

$$\det\left(\frac{1}{m} W_m W_m' - \frac{1}{\sqrt{m}} C_m + D - \frac{\phi}{\sqrt{m}} U_m\right) = 0,$$

where $W_m = \|W_{ij}^m\| : px(k-1)$, $U_m = \|U_{ij}^{(m)}\| : pxp$, with $U_{ij}^{(m)} = U_{ji}^{(m)}$,

$$D = \begin{Vmatrix} (\lambda_1-\phi)I_{\mu_1} & & & \\ & \ddots & & \\ & & (\lambda_\nu - \phi)I_{\mu_\nu} & \\ & & & -\phi I_{p-r} \end{Vmatrix}$$

$$C_m = \begin{Vmatrix} C_{ii}^{(m)},\ldots,C_{i\nu}^{(m)}; & E_1^{(m)\prime} \\ \cdots\cdots\cdots\cdots & \\ C_{\nu i}^{(m)},\ldots,C_{\nu\nu}^{(m)} & E_\nu^{(m)\prime} \\ E_{\nu 1}^{(m)},\ldots,E_\nu^{(m)} & 0 \end{Vmatrix}$$

$$C_{gh}^{(m)} = \left|\left|\sqrt{\lambda_g}\, W_{ij}^{(m)} + \sqrt{\lambda_h}\, W_{ij}^{(m)}\right|\right| , \; i = a_{h-1}+1,\ldots,a_h \; , \; j = a_{g-1}+1,$$

$$\ldots,a_g, \; 1 \leq h \leq g \leq \nu$$

$$E_h^{(m)} = \left|\left|\sqrt{\lambda_h}\, W_{ij}^{(m)}\right|\right| , \; i = r+1,\ldots,p \; , \; j = a_{h-1}+1,\ldots,a_h$$

and

$$a_0=0 \; , \; a_h = a_{h-1} + \mu_h \; , \; h = 1,2,\ldots,\nu. \quad a_\nu = r \leq P \; , \; \lambda_1 > \ldots > \lambda_\nu > 0.$$

Let $\phi_1^{(m)} \geq \ldots \geq \phi_p^{(m)} \geq 0$ be roots of this determinantal equation and let $Z_i^{(m)} = \sqrt{m}(\phi_i^{(m)} - \lambda_h)$, $i = a_{h-1}+1,\ldots,a_h$, $h = 1,2,\ldots,\nu$, and $Z_i^{(m)} = m\phi_i^{(m)}$, $i = r+1,\ldots,p$. Then the joint distrubution of $(Z_i^{(m)},\ldots,Z_p^{(m)})$ tends in distribution to that of $Z_1,\ldots,Z_p$, where $Z_{a_{h-1}+1} \geq \ldots \geq Z_{a_h}$ are the roots of

$$\det(C_{hh} + \lambda_h U_h - Z I_{\mu_h}) = 0 \; , \; h = 1,2,\ldots,\nu,$$

and $Z_{r+1} \geq \ldots \geq Z_p$ are the roots of

$$\det(\left|\left| d_{ij} \right|\right| - Z I_{p-r}) = 0,$$

where

$$C_{kk} = \left|\left| \lambda_h (W_{ij} + W_{ji}) \right|\right| \; , \; i \; , \; j = a_{h-1}+1,\ldots,a_h,$$

$$U_{kk} = \left|\left| U_{ij} \right|\right| \; , \; j = a_{h-1}+1,\ldots,a_h \; , \; U_{ij} = U_{ji}$$

$$a_{ij} = \sum_{\ell=r+1}^{k-1} W_{i\ell}\, W_{j\ell} \; , \; i, \; j = r+1,\ldots,p.$$

PROOF.

According to Theorem 1, without loss of generality, we can assume that $W_{ij}^{(m)} \to W_{ij}$ and $U_{ij}^{(m)} \to U_{ij}^{(m)} \to U_{ij}$ pointwise. The explicit proof refers to [8] and is omitted here.

4.6 (See [2], [5], [6])

Suppose that $(X_{n1},X_{n2},\ldots,X_{nk}) \xrightarrow{w} (X_1,\ldots,X_k)$ for any k, and that $\overline{\lim_{K\to\infty}}\ \overline{\lim_{n\to\infty}} \sum_{k=K}^{\infty} |X_{nk}| = 0$ in probability, $\sum_{k=1}^{\infty} X_{nk}$ and $\sum_{k=1}^{\infty} X_k$ a.s. converges. Also suppose that $g_k(t)$ is uniformly bounded in k and t. Then the sequence of stochastic processes $\sum_{k=1}^{\infty} X_{nk}g_k(t)$ weakly converges to the stochastic process $\sum_{k=1}^{\infty} X_k g_k(t)$.

PROOF.

Set $S = \{(x_1,x_2,\ldots,): \sum_{k=1}^{\infty} |x_k| < \infty\}$. Define

$\rho\left((x_1, x_2,\ldots,), (y_1, y_2,\ldots,)\right) = \sum_{k=1}^{\infty} |x_k-y_k|$. Then it is easy to see that S is a complete separable metric space and $(X_{n1}, X_{n2},\ldots)$, $(X_1, X_2,\ldots,)$ are random elements on S with property

$$(X_{n1},\ldots,X_{nk},\ldots,) \xrightarrow{w} (X_1,\ldots, X_k,\ldots,).$$

According to Theorem 1, we can assume that this convergence is true pointwise. It is not difficult to show that

$$\left|\sum_{k=1}^{\infty} X_{nk}\, g_k(t) - \sum_{k=1}^{\infty} X_k g_k(t)\right|$$
$$\leq M \sum_{k=1}^{\infty} |X_{nk}-X_k| \to 0. \quad \text{a.s.}$$

and the proof is complete.

For details of this example, the reader is referred to Bai and Yin (1984). The proof of Theorem 5.2 given there can be greatly simplified by using Theorem 1.

In all the above examples, we can use Skorokhod's Theorem. In the following we shall give an example to show that Skorokhod's Theorem is unapplicable.

4.7

Suppose that $X_p = (X_{ij})$: pxn and $T_p = (t_{ij}^{(p)})$ satisfy

1) $\{X_{ij}, i,j=1,2,\ldots\}$ are i.i.d. random variables with mean zero and variance $\sigma^2 > 0$.

2) For each p, Tp is a non-negative definite random matrix.

3) X_p is independent of Tp.

4) $\frac{1}{p}$ trace $t_p^k \xrightarrow{\text{in P}} H_k$ as $p \to \infty$, for each k.

5) $\frac{p}{n} \longrightarrow y \in (0, \infty)$, $p \to \infty$.

Then for any $k \geq 1$

$$\frac{1}{p} \text{ trace } \left(\frac{1}{n} X_p X_p' T_p\right)^k \longrightarrow E_k \text{ in pr.}$$

where E_k is a constant depending only upon σ, y and $H_1,\ldots,H_k$, (See [9]).

PROOF.

Take $S_p = R^{np + \frac{1}{2}P(P+1)}$, the Euclidean Space

$\Phi_p = \{\frac{1}{p} \text{ trace } T_p^i, i = 1,2,\ldots,k\}$: $S_p \longmapsto [0, \infty)^k$ and ν_p the measure on S_p, derived by (X_p, T_p).

By the assumptions we have

$$\mu_p \, \phi_p^{-1} \xrightarrow{w} E_{(H_1,\ldots,H_k)}.$$

Thus, we can assume, by Theorem 1, that for fixed k,

$$\{\frac{1}{p} \text{ trace } T_p^i \ , \ i = 1,2,\ldots,k\} \longrightarrow \{H_1,\ldots,H_k\} \quad \text{pointwise.}$$

After truncation and centralization on $\{X_{ij} \ , \ i \ , \ j=1,2,\ldots,\}$, we can prove that

$$E[\frac{1}{p} \text{ trace } \{\frac{1}{p} \tilde{X}_p \, \tilde{X}_p' \, T_p\}^k | T_p] \longrightarrow E_k \ , \ p \to \infty.$$

and

$$\sum_{p=1}^{\infty} E\{[\frac{1}{p} \text{ trace } \{\frac{1}{p} \tilde{X}_p \, \tilde{X}_p' \, T_p\}^k - H_k]^2 | T_p\} < \infty$$

where $\tilde{X}_p = ||X_{ij}^{(p)}||$, pxn and $X_{ij}^{(p)}$ is the random variable obtained from X_{ij} by truncation and contralization. Thus

$$\frac{1}{p} \text{ trace } \{\frac{1}{p} X_p \, X_p' \, T_p\}^k \longrightarrow E_k, \text{ a.s.}$$

and consequently

$$\frac{1}{p} \text{ trace } \{\frac{1}{p} X_p \, X_p' \, T_p\}^k \longrightarrow E_k, \text{ a.s.}$$

Since we have used Theorem 1, the above expression only implies that the convergence to be proved is true in probability. The details of the proof can refer to [9].

Note that in this example Skorohod's Theorem is unapplicable, because S_p, defined here, is not the same.

REFERENCES

[1] Bai, Z.D. (1984). A Note on the Joint Distribution of the Eigenvalues of Noncentral Multivariate F Matrix. Center for Multivariate Analysis, University of Pittsburgh. Technical Report No. 84-49.

[2] Bai, Z.D. and Yin, Y.Q. (1984). Limiting Behavior of the Norm of Products of Random Matrices and Two Problems of Geman and Hwang. Center for Multivariate Analysis, University of Pittsburgh, Technical Report No. 84-46.

[3] BiMingsley (1968). Convergence of Probability Measures. Wiley, New York.

[4] Chung, K.L. (1974). A Course in Probability Theory. Academic Press, New York.

[5] Geman, S. and Hwang, C.R. (1982). A Chaos Hypothesis for Some Large Systems of Random Equations. Z. Wahrscheinlichkeistheoria Verw. Gehiete, No. 60, 291-314.

[6] Geman, S. (1982). Almost Sure Stable Oscillations in a Large System of Randomly Coupled Equations. Siam Journal of Applied Math, Vol. 42, No.4 695-703.

[7] Hsu, P.L. (1941). On the Limiting Distribution of Roots of a Determinantal Equation. J. London Math. Soc. 16, 183-194.

[8] Bai, Z.D., Krishnaiah, P.R. and Liang, W.Q. (1984). On Asymptotic Joint Distribjtion of the Eigenvalues of the Noncentral MANOVA Matrix for Non-normal Populations. Center for Multivariate Analysis, University of Pittsburgh. Technical Report No. 84-53.

[9] Yin, Y.Q. (1984). Limiting Sepetral Distribution for a Class of Random Matrices. Center for Multivariate Analysis, University of Pittsburgh. Technical Report No. 84-50.

[10] Skorokhod, A.V. (1956). Limit Theorems for Stochastic Processes. Theory of Prob. and its Appl. Vol. 1, No.3, 261-290.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 85-0683	2. GOVT ACCESSION NO. AD-A159078	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Strong Representation of Weak Convergence	5. TYPE OF REPORT & PERIOD COVERED Technical - July 1985
	6. PERFORMING ORG. REPORT NUMBER 85-29
7. AUTHOR(s) Z. D. Bai and W. Q. Liang	8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis 515 Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Department of the Air Force Bolling Air Force Base, DC 20332	12. REPORT DATE July 1985
	13. NUMBER OF PAGES 39
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Eigenvalues, Random Matrices, Strong Representation, Weak Convergence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let μ_n, $n = 1,2,\ldots$, and μ be a given sequence of probability measures each of which is defined on a complete separable metric space S_n and S respectively. Also, a sequence of measurable mappings ϕ_n from S_n into S is given. In this paper, it is proved that if $\mu_n \circ \phi_n^{-1}$ weakly converge to μ, then there is a probability space (Ω,F,P), on which we can define a sequence

(turn over)

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(Block 20 "Abstract" Continued)

of random elements X_n, from Ω into S_n, and a random element X, from Ω into S, such that ν_n is the distribution of X_n, μ is the distribution of X and $\lim_{n\to\infty} \rho_n(X_n) = X$ pointwise.

The result of Skorokhod (1956) is a special case of the result of this paner. Some applications in the area of random matrices, etc., are also given.

END

FILMED

11-85

DTIC